THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 7, September, 1973
Printed in U.S.A.

Distribution of $\sum a_n/n$, a_n Randomly Equal to ± 1

By S. O. RICE

(Manuscript received February 6, 1973)

When $a_1, a_2, \cdots$ are independent random variables, each equal to ± 1 with probability $\frac{1}{2}$, the sum $\sum_1^\infty a_n/n$ is a random variable whose distribution is difficult to determine theoretically. This sum is of interest in the study of intersymbol interference in digital communication systems. Here the distribution of the sum is computed by numerical integration and the results tabulated. Asymptotic expressions are given for the tails of the distribution.

I. INTRODUCTION

The distribution of the random variable

$$x = \sum_{n=1}^{\infty} a_n/n, \tag{1}$$

where $a_1, a_2, \cdots$ are independent random variables equal to $+1$ or -1 with probability $\frac{1}{2}$, is of some interest in the study of intersymbol interference in a digital communication system. For example, the sum of two independent expressions of the form (1) occurs when the pulse train $\sum_{-\infty}^{\infty} a_n \sin(t - n\pi)/(t - n\pi)$, a_n randomly equal to ± 1, is sampled at regularly spaced instants which are slightly out of step with the zeros of $\sin t$. The theory of random variables of type (1) (in particular with $a_n\beta^n$ in place of a_n/n) has been studied by a number of investigators. A survey of the field has been made recently by Hill and Blanco.[1] Here we evaluate the distribution of x numerically and give expressions for its behavior when x is large. Questions of continuity and convergence are put aside.

Since the distribution is even about $x = 0$, only values for $x \geqq 0$ need be considered. From the characteristic function

$$f(u) = \text{avg}\,[\exp(ixu)] = \prod_{n=1}^{\infty} [\cos(u/n)] \tag{2}$$

we get an expression for the probability density $p(x)$ of x:

$$p(x) = \frac{1}{\pi} \int_0^\infty \cos(xu) f(u) du, \tag{3}$$

$$\text{Prob}(x > x_1) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\sin(x_1 u)}{u} f(u) du. \tag{4}$$

The values of $p(x)$ and Prob $(x > x_1)$ sbown in Table I were obtained by evaluating these integrals by the trapezoidal rule[2,3] which works well for (3) and (4).

The asymptotic expressions (8) and (18) for $p(x)$ follow from a saddle point analysis of (3). Botb $p(x)$ and Prob $(x > x_1)$ decrease rapidly wben x (or x_1) > 3, the decrease being dominated by the factor $\exp[-\exp(x - A)]$ where $A = 1.39 \cdots$.

The rapid decrease of $p(x)$ is interesting because the divergence of $\sum 1/n$ migbt lead one to expect tbat $p(x)$ would decrease slowly as $x \to \infty$. Instead, $p(x)$ actually decreases mucb faster tban a Gaussian probability density. I am indebted to a referee for the observation that tbe second, fourth, and sixth moments of $p(x)$ are, respectively, $\pi^2/6$, $11\pi^4/180$, and $233\pi^6/7560$.

Tbe reader may wonder why as many as six decimal places are given in Table I. There are several reasons. One is that tbe cost was low. About 3 seconds were required by a Honeywell 6000 Processor to compute tbe values sbown in Table I, and about 40 terms were required in each trapezoidal sum. This illustrates the fact (apparently not well known) that when integrals like (3) and (4) are to be evaluated numerically, the trapezoidal rule often performs better than most of the other conventional quadrature methods (better than Simpson's rule, for instance).[2] The six-figure accuracy is also used to gain an idea of tbe values of x for which the asymptotic expansion (18) for $p(x)$ begins to be valid. This degree of accuracy also shows that $p(0)$ is equal to $0.249\ 994 \cdots$ and not to $\frac{1}{4}$, as might be inferred from a four-figure tabulation.

II. TRAPEZOIDAL RULE CALCULATION

Preliminary computations sbowed that $|f(u)| < 10^{-10}$ when $u > 15$. Furtbermore, it was found that (3) and (4) could be evaluated to within tbe desired accuracy by using a trapezoidal-rule spacing of $\Delta u = h = 0.4$. In line witb these values the trapezoidal sum was truncated at the 40tb term $(15/0.4 \approx 40)$.

Table I—Values of $p(x)$ and Prob $(x > x_1)$

x or x_1	$p(x)$	Prob $(x > x_1)$	x or x_1	$p(x)$	Prob $(x > x_1)$
0.0	0.249 994	0.500 000	2.0	0.125 000	0.056 599
0.2	0.249 970	0.450 003	2.2	0.091 768	0.034 949
0.4	0.249 802	0.400 021	2.4	0.061 647	0.019 683
0.6	0.249 073	0.350 118	2.6	0.037 148	0.009 912
0.8	0.246 778	0.300 494	2.8	0.019 592	0.004 357
1.0	0.241 222	0.251 623	3.0	0.008 777	0.001 623
1.2	0.230 408	0.204 357	3.2	0.003 222	0.000 494
1.4	0.212 852	0.159 912	3.4	0.000 927	0.000 118
1.6	0.188 353	0.119 683	3.6	0.000 198	0.000 021
1.8	0.158 232	0.084 949	3.8	0.000 029	0.000 003
2.0	0.125 000	0.056 599	4.0	0.000 003	0.000 000

The infinite product (2) for $f(u)$ was computed by using

$$f(u) = \left(\prod_{n=1}^{N-1} [\cos (u/n)]\right) \exp\left[\sum_{n=N}^{\infty} \ln \cos (u/n)\right], \tag{5}$$

where N is a large number such that $u/N \ll 1$ for all values of u used in the computation $(0 \leqq u \leqq 16)$. The product $\prod_1^{N-1}$ in (5) was computed by straightforward multiplication. The sum in (5) was computed by setting $g(n) = \ln [\cos (u/n)]$ in the Euler-Maclaurin sum formula:

$$\sum_{n=N}^{\infty} g(n) = \int_N^{\infty} g(t)dt + \frac{1}{2} g(N) - \frac{B_2}{2\,!} g^{(1)}(N) - \frac{B_4}{4\,!} g^{(3)}(N) - \cdots - \frac{B_{2k}}{(2k)\,!} g^{(2k-1)}(N) + R_k. \tag{6}$$

Here the B's denote Bernoulli's numbers, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$ [we stopped at B_6 in our use of (6)], $g^{(k)}(N)$ denotes the value of $(d/dt)^k g(t)$ at $t = N$, and the remainder R_k is the integral of $g^{(2k+1)}(t)$ times the Bernoulli "polynomial" of degree $2k + 1$ and period 1 (see pages 520–540 of Ref. 4).

The integral in (6) can be evaluated to within the desired accuracy by setting $u/t = y$, $dt = -\,udy/y^2$, expanding $\ln (\cos y)$ in powers of y with the help of

$$-\ln (\cos y) = \int_0^y \tan v dv = \int_0^y \left[v + \frac{v^3}{3} + \frac{2v^5}{15} + \frac{17v^7}{315} + \cdots\right] dv,$$

and integrating termwise:

$$\int_N^\infty g(t)dt = u \int_0^{u/N} y^{-2} \ln(\cos y)dy$$
$$= \frac{u^2}{2N} + \frac{u^4}{36N^3} + \frac{u^6}{225N^5} + \frac{17u^8}{17640N^7} + \cdots. \tag{7}$$

Expressions for tbe higher derivatives of $g(N)$ in (6) can be ohtained by differentiating tbe series in (7) with respect to N.

In using (6) we stopped at $k = 3$ and neglected R_3.

Three separate trapezoidal-rule evaluations of $p(x)$ and Prob $(x > x_1)$ were made using eqs. (3) to (7) with $h = 0.4$, $N = 201$; $h = 0.38$, $N = 201$; and $h = 0.36$, $N = 301$, respectively. Here h is the spacing used in the trapezoidal-rule evaluations of (3) and (4). Tbe three sets of computed values differed only in tbe 7th or 8th decimal places, i.e., all agreed with the values sbown in tbe table. The values 201 and 301 of N are so large that terms beyond $k = 3$ in (6) and tbose shown in (7) are not needed. To check the computations, the integrals of $p(x)$ amd $x^2p(x)$ from $x = 0$ to $x = \infty$ (the upper limit used was actually $x = 5$) were computed by the trapezoidal rule with a spacing of $\Delta x = 0.1$. Tbe trapezoidal values agreed with the known values, respectively $\frac{1}{2}$ and $\pi^2/12$, to within 6 significant figures or better.

III. DISCUSSION OF TABLE I

Table I shows that $p(x)$ remains nearly equal to $p(0) = 0.249994$ for $0 \leqq x \leqq 1$, passes through $p(2) = 0.125000$ (is it exactly $\frac{1}{8}$?), and then decreases rapidly. The question as to whether $p(2)$ is exactly $\frac{1}{8}$ remains unanswered, but $p(0) = 0.249994$ does not seem to be an erroneous calculation of $\frac{1}{4}$. For if $p_2(u)$ is the probability density of $u = \sum_2^\infty a_n/n$, then

$$p(x) = \tfrac{1}{2}p_2(x-1) + \tfrac{1}{2}p_2(x+1).$$

Setting x equal to 0 and 2 and combining tbe results give a result I owe to J. E. Mazo,

$$p(2) = \tfrac{1}{2}p(0) + \tfrac{1}{2}p_2(3) > \tfrac{1}{2}p(0).$$

Furthermore, replacing $\frac{1}{2}p_2(3)$ hy $p(4) - \frac{1}{2}p_2(5)$ gives

$$p(2) = \tfrac{1}{2}p(0) + p(4) - \tfrac{1}{2}p_2(5)$$
$$0.125000 = 0.124997 + 0.000003 - \tfrac{1}{2}p_2(5)$$

which is satisfied by the tabulated values when $p_2(5)$ is assumed to he negligibly small.

IV. ASYMPTOTIC EXPRESSIONS FOR LARGE x

We shall show that the rapid decrease of $p(x)$ for $x > 3$ is described by

$$
\begin{aligned}
p(x) &\sim (y_0/\pi)^{\frac{1}{2}} e^{-y_0}, \\
y_0 &= \exp\left[x - 2\gamma + \ln(\pi/4)\right] = \exp\left[x - 1.39599 \cdots\right],
\end{aligned}
\tag{8}
$$

where γ denotes Euler's constant, $0.577215 \cdots$. Integrating (8) gives

$$\text{Prob}\,(x > x_1) \sim \operatorname{erfc}\,(y_0^{\frac{1}{2}}) \sim p(x_1)/y_0, \tag{9}$$

where y_0 is computed from the second of equations (8) with x_1 in place of x. Bounds for the distribution involving exponential functions of e^x have been obtained by L. A. Shepp in unpublished work.

To obtain (8) we rewrite the integral (3) for $p(x)$ as

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixu} f(u)\,du. \tag{10}$$

As is often the case for such integrals, the asymptotic value of $p(x)$ is given by the contribution of a saddle point, $u_0 = iy_0$, lying far out on the positive imaginary u-axis (the path of integration being deformed so as to pass through the saddle point). For $u = iy$ the integrand in (10) becomes

$$\exp\left[-xy + \varphi(y)\right], \qquad \varphi(y) = \sum_{n=1}^{\infty} \ln\left[\cosh\,(y/n)\right]. \tag{11}$$

When y is large we can show that

$$
\begin{aligned}
\varphi(y) &= y \ln y - y + Ay + \tfrac{1}{2} \ln 2 + r(y), \\
A &= 2\gamma - \ln(\pi/4) = 1.39599 \cdots,
\end{aligned}
\tag{12}
$$

where $r(y)$ has roughly the same magnitude as $\exp(-2y)$.

We first outline the derivation of the expression (12) for $\varphi(y)$, and then apply (12) to obtain the asymptotic expression (8) for $p(x)$.

The derivation of (12) is based upon the Euler-Maclaurin sum formula (6) with $N = 1$ and $g(t) = \ln\left[\cosh\,(y/t)\right]$. The integral in the sum formula is

$$
\begin{aligned}
&\int_1^{\infty} \ln\left[\cosh\,(y/t)\right] dt \\
&\quad = y \int_0^{y} v^{-2} \ln\,(\cosh v)\,dv \\
&\quad = -\ln\,(\cosh y) + y(\ln y) \tanh y - y \int_0^{y} (\ln v) \operatorname{sech}^2 v\,dv \\
&\quad = \ln 2 - y + y \ln y - y \int_0^{\infty} (\ln v) \operatorname{sech}^2 v\,dv + 0\left[y e^{-2y} \ln y\right],
\end{aligned}
\tag{13}
$$

where we have integrated by parts twice. The last integral in (13) has the value

$$\int_0^\infty (\ln v) \operatorname{sech}^2 v dv = -\gamma + \ln(\pi/4) \tag{14}$$

which can be obtained by (*i*) replacing $\ln v$ in (14) by $v^{\mu-1}$, (*ii*) differentiating the known value (formula 3.527-3, page 352 of Ref. 5) of the resulting integral with respect to μ, and (*iii*) setting $\mu = 1$. The derivative of $g(t)$ with respect to t, $g^{(1)}(t)$, in the sum formula (6) is

$$\begin{aligned} g^{(1)}(t) &= -yt^{-2} \tanh(y/t) \\ &= -yt^{-2}(1 - 2e^{-2y/t} + 2e^{-4y/t} - \cdots), \end{aligned}$$

where the exponential terms become negligible when y becomes large and $t = 1$. In general, for $l = 0, 1, 2, \cdots, g^{(2l+1)}(1)$ is equal to $-(2l+1)!\,y$ plus negligible terms. Therefore, the right side of the sum formula (6) is the integral plus

$$\tfrac{1}{2}(y - \ln 2) + \frac{B_2}{2} y + \frac{B_4}{4} y + \cdots + yR'_k \tag{15}$$

plus terms which are negligible when y is large. The sum of the coefficients of y in (15) is known to be equal to γ (page 529 of Ref. 4). Hence (15) is equal to $\gamma y - \frac{1}{2}\ln 2$. Addition of (13) and (15) and use of (14) gives the expression (12) for $\varphi(y)$.

Next we use the expression for $\varphi(y)$, with the small term $r(y)$ neglected, to obtain the asymptotic form of $p(x)$. The saddle point of interest occurs at $u_0 = iy_0$ where y_0 is the zero of the derivative of the exponent in (11). The exponent is $-xy + \varphi(y)$, and y_0 is the zero of

$$-x + \varphi'(y) = -x + \ln y + A.$$

Thus $y_0 = \exp(x - A)$. This y_0 is the same as the y_0 appearing in the asymptotic expression for $p(x)$ stated in (8). The exponent itself has the value $-xy_0 + \varphi(y_0) = -y_0 + \frac{1}{2}\ln 2$ at y_0. By making use of $\varphi''(y) = 1/y$ and the higher derivatives of $\varphi(y)$, the exponent can be expanded in a Taylor series about y_0. From this expansion it follows that near u_0 the integrand in the integral (10) for $p(x)$ can be written as

$$\exp\left[-y_0 + \frac{1}{2}\ln 2 + y_0 \sum_{k=2}^{\infty} \frac{(iz/y_0)^k}{(k-1)k}\right], \tag{16}$$

where $z = u - u_0$. Setting (16) in (10), changing the variable of integration from u to $\tau = z/y_0 = (u - u_0)/y_0$, and assuming that $p(x)$ is given asymptotically (as x and y_0 tend to ∞) by the contribution of

the saddle point at u_0 give

$$p(x) \sim (2\pi)^{-1}2^{\frac{1}{2}}y_0e^{-y_0}\int \exp\left[y_0\sum_{k=2}^{\infty}\frac{(i\tau)^k}{(k-1)k}\right]d\tau. \tag{17}$$

Here the nominal path of integration is the real τ-axis. The classical saddle point asymptotic expansion obtained from (17) is

$$p(x) \sim (y_0/\pi)^{\frac{1}{2}}e^{-y_0}\left[1+\frac{1}{24y_0}-\frac{23}{1152y_0^2}+\cdots\right]. \tag{18}$$

The coefficients of the powers of $1/y_0$ in the series can be determined by a general procedure described in Appendix D, page 1999, of Ref. 6.

The asymptotic expression stated in (8) is the leading term in (18). An idea of the accuracy of the asymptotic expressions can be obtained by considering the case $x = 3$. For $x = 3$, y_0 is 4.973 and Table I gives the "exact" value $p(3) = 0.008777$. The asymptotic value of $p(3)$ obtained from (8) is 0.008710, the first two terms in (18) give 0.008783, and the first three terms give 0.008776. Table I gives the "exact" value Prob $(x > 3) = 0.001623$ and eq. (9), namely Prob $(x > 3)$ $\sim$erfc $(y_0^{\frac{1}{2}})$ = erfc (2.230), gives 0.001612.

REFERENCES

1. Hill, F. S., Jr., and Blanco, M. A., "Random Geometric Series and Intersymbol Interference," accepted for publication in IEEE Trans. Information Theory.
2. Rice, S. O., "Efficient Evaluation of Integrals of Analytic Functions by the Trapezoidal Rule," B.S.T.J., *52*, No. 5 (May–June 1973), pp. 707–722.
3. Kendall, D. G., "A Summation Formula Associated With Finite Trigonometric Integrals," Quart. J. Math. (Oxford Ser.), *13* (1942), pp. 172–184.
4. Knopp, K., *Theory and Application of Infinite Series*, London: Blackie and Son, 1928.
5. Gradshteyn, I. S., and Ryzhik, I. W., *Tables of Integrals, Series, and Products*, New York and London: Academic Press, 1965.
6. Rice, S. O., "Uniform Asymptotic Expansions for Saddle Point Integrals—Application to a Probability Distribution Occurring in Noise Theory," B.S.T.J., *47*, No. 9 (November 1968), pp. 1971–2013.